



(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 885 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) INITIAL ORIGIN:

- (A) ORGANISM: hepatitis C virus

(viii) POSITION IN GENOME:

- (A) CHROMOSOME SEGMENT: NS3

(ix) CHARACTERISTICS:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..885

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATG ACC ATG ATT ACG AAT TCC CGG GGA TCC ATC ATG AAA TCC CCG GTG 48  
Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ile Met Lys Ser Pro Val  
1 5 10 15

TTC ACG GAT AAC TCC TCT CCA CCG GTA GTG CCC CAG AGC TTC CAG GTG 96  
Phe Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val  
20 25 30

GCT CAC CTG CAT GCT CCC ACA GGC AGC GGC AAG AGC ACC AAG GTC CCG 144  
Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro  
35 40 45

GCT GCA TAC GCA GCT CAG GGC TAC AAG GTG CTA GTG CTC AAC CCT TCT 192  
Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser  
50 55 60

GTT GCT GCA ACA TTG GGC TTT GGT GCC TAC ATG TCC AAG GCT CAT GGG 240  
Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly  
65 70 75 80

ATC GAT CCT AAC ATC AGG ACC GGG GTG AGA ACA ATT ACC ACT GGC AGC 288  
Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser  
85 90 95

CCC ATT ACG TAC TCC ACT TAC GGC AAG TTT CTT GCC GAC GGC GGG TGC 336  
Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys  
100 105 110

[illegible]

(3) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 295 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ile Met Lys Ser Pro Val  
1 5 10 15  
Phe Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val  
20 25 30  
Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro  
35 40 45  
Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser  
50 55 60  
Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly  
65 70 75 80  
Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser  
85 90 95  
Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys  
100 105 110  
Ala Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser Thr  
115 120 125  
Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Gly Glu  
130 135 140  
Thr Ala Gly Ala Lys Leu Val Val Phe Ala Thr Ala Thr Pro Pro Gly  
145 150 155 160  
Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr  
165 170 175  
Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile  
180 185 190  
Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Arg Lys Cys Asp  
195 200 205

Glu Leu Ala Thr Lys Leu Val Ala Met Gly Ile Asn Ala Val Ala Tyr  
 210 215 220  
 Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val Val  
 225 230 235 240  
 Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp  
 245 250 255  
 Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe Ser  
 260 265 270  
 Leu Asp Pro Thr Phe Thr Ile Glu Thr Thr Thr Leu Pro Gln Asp Ala  
 275 280 285  
 Val Ser Arg Thr Gln Arg Arg  
 290 295

(4) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 40 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AAGGGATCCA TCATGAAATC CCCGGTGTTC ACGGATAACT

40

(5) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GGGAAGCCTT AATTCTTACC GTCGTTGAGT GCGGGAGAC

(6) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GAGGGATCCA TCATGAAAGC GGTGGACTTT ATCCCTGTG

39

(7) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 33 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GAGAAGCTTT TAACACGTGT TGCAGTCTAT CAC

33

(8) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GAGGGATCCA TCATGAAACA CCTGCATGCT CCCACCGGC

39

1059902609660

(8) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 33 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GAGAAGCTTT TAATACCAAG CACAGCCTGC GTC

33

(3) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 302 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Pro Val Glu Asn Leu Glu Thr Thr Met Arg Ser Pro Val Phe Thr  
1 5 10 15  
Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val Ala  
20 25 30  
His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro  
35 40 45  
Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro  
50 55 60  
Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala  
65 70 75  
His Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr  
80 85 90  
Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala  
95 100 105

13.

assum 2

This is  
14-forward  
in 2

[illegible]





(vi) INITIAL ORIGIN:

(A) ORGANISM: hepatitis C virus

(viii) POSITION IN THE GENOME:

(A) CHROMOSOME/SEGMENT: NS3

(ix) CHARACTERISTICS:

(A) NAME/KEY: CDS

(B) LOCATION: 1..885

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATG ACC ATG ATT ACG AAT TCC CGG GGA TCC ATC ATG AAA TCC CCG GTG	48
Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ile Met Lys Ser Pro Val	
1 5 10 15	
TTC ACG GAT AAC TCC TCT CCA CCG GTA GTG CCC CAG AGC TTC CAG GTG	96
Phe Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val	
20 25 30	
GCT CAC CTG CAT GCT CCC ACA GGC AGC GGC AAG AGC ACC AAG GTC CCG	144
Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro	
35 40 45	
GCT GCA TAC GCA GCT CAG GGC TAC AAG GTG CTA GTG CTC AAC CCT TCT	192
Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser	
50 55 60	
GTT GCT GCA ACA TTG GGC TTT GGT GCC TAC ATG TCC AAG GCT CAT GGG	240
Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly	
65 70 75 80	
ATC GAT CCT AAC ATC AGG ACC GGG GTG AGA ACA ATT ACC ACT GGC AGC	288
Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser	
85 90 95	
CCC ATT ACG TAC TCC ACT TAC GGC AAG TTT CTT GCC GAC GGC GGG TGC	336
Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys	
100 105 110	
GCA GGG GGT GCT TAT GAC ATA ATA ATT TGT GAC GAG TGC CAC TCC ACG	384
Ala Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser Thr	
115 120 125	
GAT GCC ACA TCC ATC TTG GGC ATC GGC ACT GTC CTT GAC CAA GGA GAG	432
Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Gly Glu	
130 135 140	

ACT Thr 145	GCG Ala	GGG Gly	GCG Ala	AAA Lys	TTG Leu 150	GTT Val	GTG Val	TTC Phe	GCC Ala	ACC Thr 155	GCC Ala	ACC Thr	CCT Pro	CCG Pro	GGC Gly 160	480
TCC Ser	GTC Val	ACT Thr	GTG Val	CCC Pro 165	CAT His	CCC Pro	AAC Asn	ATT Ile	GAG Glu 170	GAG Glu	GTT Val	GCT Ala	CTA Leu	TCC Ser 175	ACC Thr	528
ACC Thr	GGA Gly	GAG Glu	ATC Ile 180	CCT Pro	TTT Phe	TAC Tyr	GGC Gly	AAG Lys 185	GCT Ala	ATC Ile	CCC Pro	CTT Leu	GAG Glu 190	GTA Val	ATC Ile	576
AAG Lys	GGG Gly	GGG Gly	AGA Arg 195	CAT His	CTC Leu	ATC Ile	TTC Phe 200	TGT Cys	CAT His	TCA Ser	AAG Lys	AGG Arg 205	AAG Lys	TGC Cys	GAT Asp	624
GAG Glu 210	CTC Leu	GCC Ala	ACA Thr	AAG Lys	CTG Leu	GTC Val 215	GCA Ala	ATG Met	GGC Gly	ATC Ile	AAT Asn 220	GCC Ala	GTG Val	GCC Ala	TAC Tyr	672
TAC Tyr 225	CGC Arg	GGT Gly	CTT Leu	GAC Asp	GTG Val 230	TCC Ser	GTC Val	ATC Ile	CCG Pro	ACC Thr 235	AGC Ser	GGT Gly	GAT Asp	GTT Val	GTC Val 240	720
GTC Val	GTG Val	GCA Ala	ACC Thr	GAC Asp 245	GCC Ala	CTC Leu	ATG Met	ACC Thr	GGC Gly 250	TAT Tyr	ACC Thr	GGC Gly	GAC Asp	TTC Phe 255	GAC Asp	768
TGG Ser	GTG Val	ATA Ile	GAC Asp 260	TGC Cys	AAC Asn	ACG Thr	TGT Cys	GTC Val 265	ACT Thr	CAG Gln	ACA Thr	GTC Val	GAT Asp 270	TTC Phe	AGC Ser	816
CTT Leu	GAC Asp	CCT Pro 275	ACC Thr	TTC Phe	ACC Thr	ATT Ile	GAG Glu 280	ACG Thr	ACC Thr	ACA Thr	CTT Leu	CCC Pro 285	CAG Gln	GAT Asp	GCT Ala	864
GTC Val	TCC Ser 290	CGC Arg	ACT Thr	CAA Gln	CGA Arg	CGG Arg 295										885

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 295 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ile Met Lys Ser Pro Val



(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRAND FORM: single
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AAGGGATCCA TCATGAAATC CCCGGTGTTT ACGGATAACT

40

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRAND FORM: single
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4

GGGAAGCCTT AATTCTTACC GTCGTTGAGT GCGGGAGAC

39

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRAND FORM: single



[illegible]

39

(i) SEQUENCE CHARACTERISTICS:

(B) TYPE: nucleic acid

(C) STRAND FORM: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8

33